REMARKS

In view of the following discussion, the Applicant submits that none of the claims now pending in the application is anticipated under the provisions of 35 U.S.C. § 102 or made obvious under the provisions of 35 U.S.C. § 103. Thus, the Applicant believes that all of these claims are now in allowable form.

I. REJECTION OF CLAIMS 1-4, 7-10 AND 13-16 UNDER 35 U.S.C. § 102

The Examiner has rejected claims 1-4, 7-10 and 13-16 in the Office Action as being anticipated by the Yamaguchi et al. patent (U.S. patent 6,026,359, issued on February 15, 2000, hereinafter Yamaguchi). The Applicant respectfully traverses the rejection.

Yamaguchi teaches a method for modifying a language model based on a change in a recognition parameter occurring between the training of the language model and the time of recognition. For example, in order to recognize speech in an input audio signal containing background noise, the original language model is trained using an arbitrary, prerecorded (stored) noise model that is combined with a stored "clean" (*e.g.*, free of background noise) speech model to form a composite noisy speech model. Jacobian matrices are then calculated from the stored noise model and the composite noisy speech model. Thus, when noisy speech in an input audio signal does not "match" the pre-existing composite noisy speech model, the composite noisy speech model is updated to form a modified noisy speech model based on a Taylor expansion using the Jacobian matrices and a difference between extracted noise from the input audio signal and the stored noise model. This modified noisy speech model is then used to process (*e.g.*, recognize speech in) the input audio signal. Yamaguchi does not teach, show or suggest, however, that the noisy speech model is derived directly from a clean speech model and a noise model in accordance with a signal-to-noise ratio.

The Examiner's attention is directed to the fact that Yamaguchi fails to disclose or suggest the novel method of recognizing speech in a noisy environment wherein a clean speech model and a noise model are interpolated <u>based on a signal-to-noise ratio</u>

to produce a noisy speech model, as claimed in Applicant's independent claims 1, 7 and 13. Specifically, Applicant's claims 1, 7 and 13, as amended, positively recite:

1. Method for performing speech recognition on an input audio signal having a speech component and a noise component, said method comprising the steps of:
   (a) obtaining at least one clean speech model;
   (b) obtaining at least one noise model;
   (c) deriving at least one noisy speech model directly from said at least one clean speech model and said at least one noise model <u>in accordance with a signal-to-noise ratio</u>; and
   (d) applying said at least one noisy speech model to extract a recognized text from the input audio signal. (Emphasis added)

7. Apparatus for performing speech recognition on an input audio signal having a speech component and a noise component, said apparatus comprising:
   means for obtaining at least one clean speech model;
   means for obtaining at least one noise model;
   means for deriving at least one noisy speech model directly from said at least one clean speech model and said at least one noise model <u>in accordance with a signal-to-noise ratio</u>; and
   means for applying said at least one noisy speech model to extract a recognized text from the input audio signal. (Emphasis added)

13. A computer-readable medium having stored thereon a plurality of instructions, the plurality of instructions including instructions which, when executed by a processor, cause the processor to perform the steps of a method for performing speech recognition on an input audio signal having a speech component and a noise component, said method comprising the steps of:
   (a) obtaining at least one clean speech model;
   (b) obtaining at least one noise model;
   (c) deriving at least one noisy speech model directly from said at least one clean speech model and said at least one noise model <u>in accordance with a signal-to-noise ratio</u>; and
   (d) applying said at least one noisy speech model to extract a recognized text from the input audio signal. (Emphasis added)

Applicant's invention is directed to a method and apparatus for recognizing speech in a noisy environment. The abilities of conventional speech recognition systems to accurately recognize speech are often limited by the presence of

background noise at the time of speech input, which compromises the clarity of the input audio signals. Although various noise compensation schemes, such as Parallel Model Combination (PMC), have been proposed, these schemes are typically computationally intensive and require large amounts of memory. Thus, such schemes are not practical for implementation in real-time applications, which require substantially instantaneous recognition, or in portable applications, which typically have access to limited memory and processing resources.

The present invention provides a method and apparatus for recognizing speech in a noisy environment in which an acoustic model representing noisy speech is applied to the input noisy speech signal to achieve recognition. In one embodiment, the method derives the noisy speech model by interpolating between a clean speech model and a noise model to produce a noisy speech model. The noise model that is used to produce the noisy speech model is derived by extracting noise directly from the input noisy speech signal. This derived noise model is also used to determine (*e.g.*, based on an estimated signal-to-noise ratio in the input noisy speech signal) an interpolation weight to be applied in the interpolating stage (*e.g.*, a ratio in which the clean speech model and noise model should be combined). By deriving the noise model directly from the input noisy speech and using the signal-to-noise data from the input noisy speech to guide interpolation, the method is able to achieve accurate recognition in substantially fewer computational cycles than conventional speech recognition methods.

In contrast, Yamaguchi teaches a method for recognizing speech in which a difference between noise in the input speech signal and noise in a pre-existing noisy speech model is used to modify the pre-existing noisy speech model. Thus, Yamaguchi fails to anticipate or make obvious Applicant's invention.

Specifically, Yamaguchi teaches that an input speech signal containing background noise is compared to a pre-existing noisy speech model (*e.g.*, trained using arbitrary or pre-recognition background noise). Noise in the input speech signal is addressed by calculating a difference between the noise in the input speech signal and the pre-existing noise model. Yamaguchi thus fails to teach or make obvious a method of recognizing speech in a noisy environment wherein a noisy speech model used in the

recognition is derived from a noise model (which is in turn derived directly from the signal containing the speech to be recognized) and a clean speech model combined in accordance with a signal-to-noise ratio, as positively claimed by the Applicant in amended claims 1, 7 and 13.

The Examiner submits that Yamaguchi does in fact teach deriving a noisy speech model in accordance with a signal-to-noise ratio, and cites a specific passage from Yamaguchi to teach this limitation (*i.e.*, column 14, line 30 to column 15, line 5). However, the Applicant respectfully submits that the Examiner's interpretation of this passage of Yamaguchi is incorrect. The portion of Yamaguchi that the Examiner cites merely discusses the results of experiments using Yamaguchi's method for acoustic model adaptation. Signal-to-noise ratio is mentioned, in this instance, as a characteristic of the evaluation data, representing the experimental conditions under which the evaluation data were captured (*See*, column 15, lines 2-5: "The S/N ratio with respect to the evaluation data was 10db ...").

The only other reference to a signal-to-noise ratio in Yamaguchi occurs at column 12, lines 17-19 ("... and the S/N ratio of the input data is improved by subtracting the calculated average spectrum from the input data spectrum"). This passage of Yamaguchi merely recites a beneficial result of applying the spectral subtraction method.

Neither reference to signal-to-noise ratio in Yamaguchi teaches or even suggests that a signal-to-noise ratio plays any factor in the derivation of the noisy speech model (*e.g.*, by defining an interpolation weight with respect to a clean speech model and a noise model), as claimed by the Applicant. Therefore, the Applicant submits that independent claims 1, 7 and 13 fully satisfy the requirements of 35 U.S.C. §102 and are patentable thereunder.

Dependent claims 2-4, 8-10 and 14-16 depend respectively from claims 1, 7 and 13, and recite additional features therefore. As such, and for at least the reasons set forth above, the Applicant submits that claims 2-4, 8-10 and 14-16 are not anticipated by the teachings of Yamaguchi. Therefore, the Applicant submits that dependent claims 2-4, 8-10 and 14-16 also fully satisfy the requirements of 35 U.S.C. §102 and are

patentable thereunder.

II. REJECTION OF CLAIMS 5-6, 11-12 AND 17-18 UNDER 35 U.S.C. § 103

The Examiner rejected claims 5-6, 11-12 and 17-18 under 35 U.S.C. §103(a) as being unpatentable over Yamaguchi in view the Komori et al. patent (U.S. Patent No. 5,956,679, issued September 21, 1999, hereinafter Komori). The Applicant respectfully traverses the rejection.

Yamaguchi has been discussed above. Komori teaches a speech processing device that performs high-speed noise adaptation using a noise-adaptive Parallel Model Combination (PMC) model. The device extracts a non-speech interval from an input speech signal and uses data from this non-speech interval to produce a noise model. This noise model is then combined with a clean speech model in accordance with a PMC conversion to produce a noisy speech model.

The Examiner's attention is directed to the fact that Yamaguchi and Komori (either singly or in any permissible combination) fail to disclose or suggest the novel method of recognizing speech in a noisy environment wherein a clean speech model and a noise model are interpolated <u>based on a signal-to-noise ratio</u> to produce a noisy speech model, as claimed in Applicant's independent claims 1, 7 and 13. Applicant's independent claims 1, 7, and 13 have been recited above.

As recited in the preceding claim, Applicant's invention teaches a method and apparatus for recognizing speech in a noisy environment using a noisy speech model that is generated by interpolating between a clean speech model and a noise model in accordance with a signal-to-noise ratio. By deriving the noise model directly in accordance with the signal-to-noise ratio, the computational cycles normally associated with recognition of noisy speech are significantly reduced.

In contrast, neither Yamaguchi nor Komori teaches or suggests this novel approach. In fact, there is <u>no mention</u> in either Yamaguchi or Komori of using a signal-to-noise ratio to guide derivation of the noisy speech model.

Moreover, there is no suggestion or motivation to combine Yamaguchi and Komori in a manner that would yield the claimed invention. As described above, Komori

teaches that a PMC method is used to produce noisy speech models for use in speech recognition. Yamaguchi, however, teaches that PMC methods are not ideal for real-time speech recognition applications because they allegedly consume a great deal of time in training noise models and in generating noisy speech models (*See*, Yamaguchi, column 1, line 53 – column 2, line 16). Yamaguchi therefore actually teaches away from combination with Komori. Thus, the Applicant respectfully submits that the Examiner is clearly using hindsight to pick and choose elements from the references to support the rejection.

It is impermissible to use the claims as a framework from which to choose among individual references to recreate the claimed invention. *W. L. Gore Associates, Inc. v. Garlock, Inc.*, 220 U.S.P.Q. 303, 312 (1983). Moreover, the mere fact that a prior art structure could be modified to produce the claimed invention would not have made the modification obvious unless the prior art suggested the desirability of the modification. *In re Fritch*, 23 U.S.P.Q. 2d 1780, 1783, Fed. Cir. (1992); *In re Gordon*, 221 U.S.P.Q. 1125, 1127, Fed. Cir. (1984) (emphasis added). The rules applicable for combining references provide that there must be a suggestion from within the references to make the combination. *Uniroyal v. Rudkin-Wiley*, 5 U.S.P.Q. 2d 1434, 1438 (Fed. Cir. 1988); *In re Fine*, 5 U.S.P.Q. 2d at 1599 (emphasis added). Therefore, the teachings of Yamaguchi do not provide any justification for combination with the PMC methodology of Komori. Thus, independent claims 1, 7 and 13 are not made obvious by the teaching of Yamaguchi in view of Komori.

Thus, Yamaguchi and Komori fail to disclose or suggest a method recognizing speech in a noisy environment wherein a clean speech model and a noise model are interpolated based on a signal-to-noise ratio to produce a noisy speech model, for example in order to reduce computational cycles for processing an input audio signal, as claimed by the Applicant in independent claims 1, 7 and 13.

Dependent claims 5-6, 11-12 and 17-18 depend, either directly or indirectly, from claims 1, 7 and 13 and recite additional features thereof. As such and for at least the same reasons set forth above, the Applicant submits that claims 5-6, 11-12 and 17-18 are also not made obvious by the teachings of Yamaguchi in view of Komori.

Therefore, the Applicant submits that dependent claims 5-6, 11-12 and 17-18 also fully satisfy the requirements of 35 U.S.C. § 103 and are patentable thereunder.

III. CONCLUSION

Thus, the Applicant submits that all of the presented claims now fully satisfy the requirements of 35 U.S.C. §102 and §103. Consequently, the Applicant believes that all of these claims are presently in condition for allowance. Accordingly, both reconsideration of this application and its swift passage to issue are earnestly solicited.

If, however, the Examiner believes that there are any unresolved issues requiring the issuance of a final action in any of the claims now pending in the application, it is requested that the Examiner telephone Mr. Kin-Wah Tong, Esq. at (732) 530-9404 so that appropriate arrangements can be made for resolving such issues as expeditiously as possible.

Respectfully submitted,

6/28/05
Date

Kin-Wah Tong, Attorney
Reg. No. 39,400
(732) 530-9404

Moser, Patterson & Sheridan, LLP
595 Shrewsbury Avenue
Shrewsbury, New Jersey 07702